(2)

MENTAL LAPSES AND EVENT-RELATED POTENTIALS

R.R. STANNY

Naval Aerospace Medical Research Laboratory
Naval Air Station
Pensacola, Florida 32508-5700
Approved for public release; distribution unlimited.

Reviewed and approved 2 Nov 89

J. A. BRADY, CAPT, MSC USN
Commanding Officer

This research was sponsored by the Office of Naval Technology under work unit 62234N RS34H21.01 and monitored by the Navy Personnel Research and Development Center. Additional funding was provided by the Naval Medical Research and Development Command under work unit 62233N MM33P30.005.

Volunteer subjects were recruited, evaluated, and employed in accordance with the procedures specified in Department of Defense Directive 3216.2 and Secretary of the Navy Instruction 3900.39 series. These instructions are based upon voluntary informed consent and meet or exceed the provisions of prevailing national and international guidelines.

Trade names of materials and/or products of commercial or nongovernment organizations are cited as needed for precision. These citations do not constitute official endorsement or approval of the use of such commercial materials and/or products.

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NAMRL-1347	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Aerospace Medical Research Laboratory 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Navy Personnel Research and Development Center
6c. ADDRESS (City, State, and ZIP Code) Naval Air Station Pensacola, FL 32508-5700	7b. ADDRESS (City, State, and ZIP Code) San Diego, CA 92152-6800
8a. NAME OF FUNDING / SPONSORING ORGANIZATION ONT 8b. OFFICE SYMBOL (If applicable) 222	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62234N PROJECT NO. RS34H21 TASK NO. 01 WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

(U) Mental Lapses and Event-Related Potentials

12. PERSONAL AUTHOR(S)

R. R. Stanny

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim	FROM 1988 TO 1989	89-11-02	36

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Performance assessment, event-related potentials, evoked potentials, psychophysiology. (STDS)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Event-related potentials (ERPs) were recorded from 10 subjects as they performed an acoustic target detection task. Each subject's ERPs were sorted by performance level, on a trial-by-trial basis. A number of ERP measuring procedures then were compared in terms of their abilities to detect episodes of low behavioral performance. The ERP feature that proved most sensitive to fluctuations in performance was P3 (P300) amplitude. Weighted, time-averaged amplitude measures outperformed peak amplitude measures by a substantial margin. The results are discussed as they bear on using ERPs to detect lapses of equipment operator performance. Keywords:

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL J.A. BRADY, CAPT, MSC, USN, Commanding Officer	22b. TELEPHONE (Include Area Code) (904) 452-3286 22c. OFFICE SYMBOL 00

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

SUMMARY PAGE

THE PROBLEM

This report describes a study of temporal variation in human performance and associated changes in event-related potentials (ERPs). Its purpose was to compare the efficiencies of several ERP-based procedures in detecting lapses of perception and judgment. Several rapid methods of estimating mental state from ERP data were examined.

FINDINGS

All of the procedures examined discriminated reliably between ERPs recorded during episodes of low performance and ERPs recorded during periods of high performance. The best predictors were measurements of P3 amplitude. The best single predictor was a weighted time-average of P3 amplitude (the RMS-s statistic of Trejo et al (1)). This statistic outperformed conventional peak amplitude and latency measurements, regardless of whether the latter had been taken from signal averages or single trial responses. The most efficient electrode montage examined was a three-electrode array comprising frontal, central, and parietal midline sites.

RECOMMENDATIONS

Future work should be directed toward separating the responses of simultaneously active populations of neurons. These are confounded in surface measurements and are a major source of systematic measurement error. There is no valid way to distinguish between simultaneous responses without knowing their generators' locations. Therefore, it is important to develop a body of theory detailed enough to relate changes in function to changes in anatomically localizable structures.

Examining performance under conditions known to yield high rates of performance failure would be useful. In some of these conditions, ERPs differ qualitatively from ERPs recorded in alert subjects. The period of transition into sleep is an example. Incorporating the peculiarities of those conditions into existing theory should help improve the detection of performance failures.

Acknowledgments

I gratefully acknowledge the support of the late CAPT J. O. Houghton, MC, USN. Mr. S. LaCour provided able computer programming assistance. HM2 S. Eagles and Ensigns R. Holman, L. Taylor, J. Hibbler, and J. Eversole provided important assistance in the laboratory.

INTRODUCTION

The experiment discussed here was a study of some techniques that might be used to detect failures of perception and judgment in persons responsible for operating man-machine systems. Its major purpose was to compare some estimates of mental state that can be derived from short term recordings of event related potentials (ERPs).

Donchin et al. (2) have suggested that in some circumstances ERPs might be used to evaluate the workload imposed on a human operator by a particular piece of equipment. The conditions under which ERPs might be useful occur when one is interested in measuring an aspect of cognition that cannot be readily assessed by observing overt behavior. The present application affords an example: Equipment operators must work at times under conditions that produce boredom, fatigue, and other forms of stress. The problem in this context is to determine when an operator's mental resources have declined to the point at which they have become insufficient to the task at hand.

One might suppose that individuals can judge their performance by introspection. That idea has considerable face validity, but it may not be generally true. As Yeh and Wickens note (3), the empirical relation between performance and subjective workload ("effort") appears to vary according to the mental resources demanded by specific tasks. Operators undoubtedly learn the approximate relationships between effort and performance as they learn their tasks. Stress, however, can dissociate effort and resources; night shift paralysis (4) is a striking example of this phenomenon. According to resource theory (5), dissociations of effort and resources should be expected to produce dissociations of effort and performance. This is because, other factors held constant, performance is determined by the quantity and quality of the mental resources invested in a task. Thus, when operators must gauge their performance by invested effort (as when external feedback is unavailable or does not register), stress is likely to render self appraisals unreliable.

Event-related potentials are aggregate electrical fields produced when neurons discharge simultaneously in response to sensory or internally generated events. Populations of neurons that are of adequate size and appropriate geometry produce ERPs that propagate to the skin surface where they can be recorded with suitably placed electrodes (6). The locations of the cells that produce the responses visible in surface recordings are uncertain in various degrees; however, a number of correlative relationships between their discharge properties and cognition are very well established (see references 7 and 8 for reviews).

The ERP phenomenon of major a priori interest in this study was the P3 (or P300) wave of of the ERP (9). This is a comparatively large, positive going wave that occurs some 300-800 ms following a stimulus that requires a response decision. Surface recordings of P3 are generally thought to be mixtures of several different waves that occur in this range of latencies (10-12). These waves are referred to, collectively, as the 'Late Positive Complex' (LPC), a more neutral term than 'P3,' theoretically. The term 'LPC' is used in this report where a degree of theoretical neutrality seems warranted.

The present study is an analysis of ERPs obtained in a study of auditory selective attention. Observers in the study were asked to discriminate acoustic target events from nontargets. Their trial-by-trial discriminations were sorted first by response correctness and then by reaction time (RT). Then the ERPs associated with trials at different performance levels, as defined by RT, were measured by several different techniques, as will be discussed presently. Finally, the different measurement techniques were compared in terms of their abilities to distinguish between performance levels. This approach was taken under the assumption that techniques that discriminate well among mental states in the present context will prove efficient in other contexts as well.

METHODS

SUBJECTS

Twelve male volunteers participated in the experiment. The data reported here were drawn from the no-drug condition of a study of the effects of scopolamine on attention. Subjects had been given an inactive placebo before the condition in which the present data were obtained. One-half of the subjects had performed the experimental task once before, 7 d prior. The subjects' ages ranged from 21 to 24 yrs and averaged 21.9 yrs. One subject was eliminated from the analysis for a violation of experimental protocol; another's data were eliminated due to electroencephalogram (EEG) recording difficulties.

TASK PROTOCOL

Event-related potentials were recorded in an auditory selective attention paradigm. The subject was asked to respond as quickly as possible to specified target stimuli by pressing a key with his right hand.

The target stimuli were 1000-Hz tone bursts delivered to the right ear. The targets occurred in randomized sequences of 1000- and 1200-Hz tone bursts. A random 20% of all tones were 1000 Hz; 80% were 1200 Hz. A random 50% of the tones of each frequency were delivered to the left ear; the remainder were delivered to the right. The randomization of frequencies was constrained so that not more than three 1000-Hz tones were delivered in a row.

The stimuli were presented in four 9-min blocks of 272 trials each. The first 12 stimuli in each block were presented in systematically repeating order to help fix the frequencies in memory. The next 10 trials were ignored to allow time for the subjects to settle into the randomized sequences.

The stimuli were sine waves multiplied by trapezoids to produce 50-ms bursts with 9-ms onset and offset ramps. They were synthesized by a computer, converted to analog signals with 16-bit resolution at 50 kHz, attenuated to yield levels of 65 dBA, and delivered to the subjects via headphones.

TREATMENT OF THE BEHAVIORAL DATA

A correct target detection was defined as a keypress with an onset not earlier than 72 ms following the onset of a target and not later than 1000 ms following the onset of a target. A miss was scored when no keypress onset occurred during the same interval. A correct rejection was scored when no keypress onset occurred from 72 to 1000 ms relative to the onset of a nontarget. A false alarm was scored when the onset of a keypress occurred during a similar interval relative to a nontarget. Trials containing responses whose onsets occurred between -200 and +72 ms relative to stimulus onset were excluded from the analysis.

EVENT-RELATED POTENTIAL RECORDING AND ANALYSIS

General Plan of the Analysis

The general plan consisted of identifying those trials corresponding to the lower and upper quartiles of each subject's RT distribution, the slowest and fastest 25% of each subject's behavioral responses (13-15). The ERPs corresponding to the trials in RT quartile 1 were considered high-performance ERPs. Those corresponding to the trials in RT quartile 4 were considered low-performance ERPs. The different measurement techniques were then compared with respect to their abilities to distinguish the ERPs corresponding to high- and low-performance trials.

Recording Procedures

Event-related potentials were recorded from 28 EEG electrodes. The reference electrode was placed on the nose; a forehead electrode served as ground. Eye movements were monitored using vertically and horizontally arranged pairs of electro-oculogram (EOG) electrodes. The analyses presented here were based on subsets of the 28 recording sites.

The EEG and EOG were amplified, filtered between 0.1 and 100 Hz, and digitized at 250 Hz for 1200 ms beginning 200 ms before stimulus onset. Each amplifier was calibrated with 0- and 10-Hz sine waves before and after each subject's run. The amplified power of the 0-Hz sine wave was subtracted from that of the 10-Hz sine wave before calculating an amplifier's gain factor to correct for system noise.

The ERP data from individual trials were machine examined for electrical artifacts using a multistage artifact compensation and rejection protocol. The routines comprising the protocol checked each epoch and channel for motion artifacts, eye blinks, eye movements, voltage transients (electrode pops), amplifier saturation, and dead amplifiers. Ocular artifacts were compensated, when possible, using a computerized, ocular artifact filter based on the method of Gratton et al. (16). All ERP summary measurements were obtained automatically by computer programs.

ERP Waveform Summary Measurements

The following are the ERP summary measurements used in the analyses to be discussed:

(1) LPC Peak Latencies in Signal Averages. The latency of the LPC peak was taken as the length of the interval from stimulus onset to the largest positive peak between 248 and 848 ms poststimulus. Signal averages were calculated using 20 trials per performance level per subject.

(2) LPC Peak Latencies in Single Trials. In this case, each trial's ERP was first low-pass filtered with a zero-phase-shift filter yielding half power at approximately 7.5 Hz. Then, the latency of the LPC peak was measured as the length of the interval from stimulus onset to the greatest positivity from 248 to 848 ms poststimulus. Those LPC peaks whose estimated latencies fell exactly at the boundaries of the time window (248 and 848 ms poststimulus) were considered to be poorly measured and were excluded from further analysis.

(3) LPC Peak Amplitudes in Signal Averages. After estimating the latency of an LPC peak in a signal average, as described previously, its amplitude was measured as the voltage at the time corresponding to peak latency.

(4) LPC Peak Amplitudes in Single Trials. After estimating the latency of an LPC peak in a single trial, as described previously, its amplitude was measured as the voltage (before low-pass filtering) at the time corresponding to peak latency.

(5) Average amplitude of the LPC. The LPC amplitude measurements were also obtained by calculating the average amplitudes of single trial ERPs. These average values were calculated by finding the mean of the voltages recorded during the interval spanning 300-500 ms poststimulus.

(6) The RMS-s statistic of Trejo et al. (1). This is the root-mean-squared amplitude of the estimated signal-to-noise ratio of an ERP segment. To calculate the RMS-s statistic, one first calculates the conventional signal average of a set of n single-trial responses. Then each point in the signal average is divided by the standard deviation of the n voltages recorded at the corresponding poststimulus time. The result of this step is treated as a vector of signal-to-noise ratios (SNs). Finally, one produces the RMS-s statistic by calculating values of the root-mean-squared amplitude (across time) of the SN vector.

We created four SN vectors from each subject's high performance responses and four from each subject's low performance responses. Each SN vector was based on five trials, which is similar to Trejo et al., who used six trials per SN vector.

We subsequently calculated nine values of RMS-s per SN vector, one for each poststimulus time interval in a set of approximately 50-ms intervals spanning 50-500 ms poststimulus. (The precise boundaries of the intervals are shown in Tables 7 through 9.) These intervals approximate the eight, 50-ms analysis windows used by Trejo and colleagues.

Comparisons Among Electrode Montages

The montage for the first analysis consisted of electrodes placed at midline frontal, central, and parietal recording sites (International 10/20 Fz, Cz, and Pz) referred to the nose electrode. This is a widely used

array. The major ERP summary measurements considered in this analysis were LPC latency, LPC amplitude, and the nine RMS-s measurements.

A second montage analysis was performed using a set of bipolar recordings. Electrodes originally referred to nose were digitally converted to bipolar by subtracting the responses obtained from one electrode from those obtained from another. These subtractions were performed on a trial-by-trial basis. The bipolar derivations used in the analysis were F3-P3, F4-P4, F3-P4, F4-P3, O1-P3, and O2-P4. The summary measurement used in this analysis was the RMS-s statistic.

The third montage examined was a lateral array comprising symmetrically placed frontal, temporal, parietal, and occipital recording electrodes (F3, F4, T3, T4, P3, P4, O1, and O2). The main purpose in examining this montage was to compare its performance with that of the smaller Fz, Cz, Pz montage. A secondary analysis of these data afforded a comparison of vertex and nose references.

Statistical Analyses

To maintain reasonable levels of statistical power, the overall analysis was divided into several subanalyses. The division amounted to organizing the statistical tests into logical families corresponding to the different combinations of measurement type and electrode montage examined.

Repeated measures analyses of variance were performed using the BMDP 4V 'General Univariate and Multivariate Analysis of Variance' program, which reports Greenhouse-Geisser corrections for covariance matrix nonsphericity (17). The F-ratios reported here for which Greenhouse-Geisser corrections are appropriate are indicated by denoting their corrected degrees of freedom as "G-G df." Regression analyses were carried out in double precision using BMDP 9R 'All Subsets Regression.'

RESULTS AND DISCUSSION

BEHAVIORAL PERFORMANCE

A breakdown of the performance statistics is shown in Table 1, which is the pooled, stimulus-response confusion matrix. The value in each of the Table's cells is the proportion of stimuli of the type indicated by the cell's row yielding responses of the type indicated by the cell's column. The column labeled 'Invalid' contains the rates at which subjects pressed the response key during the interval between -200 to +72 ms relative to stimulus onset. Recall that target stimuli were 1000-Hz tones delivered to the left ear and that "No" responses were, in fact, nonresponses.

Responding was very accurate. The overall proportion of correct responses (the sum of correct detections and correct rejections divided by total trials) was .993. The hit rate, from Table 1, was approximately .975. The correct rejection rate, collapsed across nontargets, was .995. The proportion of targets missed was .022.

TABLE 1. The Overall Stimulus-response Confusion Matrix. Each Entry is the Proportion of the Indicated Stimulus Type Yielding the Indicated Response.

Stimulus type		Response type		
Ear	Freq	Yes	No	Invalid
Left	1200	.001	.997	.002
	1000	.005	.994	.001
Right	1200	.004	.995	.001
	1000	.975	.022	.002

False-alarm rates averaged only .003 across the three categories of nontarget. False alarms to nontargets that shared one or the other of the target's distinguishing features (frequency or ear of delivery) were somewhat more frequent than false alarms to the nontarget that shared none.

Figure 1 shows the distribution of reaction times to correctly detected targets. The mean of the RT distribution was 435 ms; its standard deviation was 107 ms. The distribution is positively skewed and of conventional appearance.

Table 2 shows the mean RTs for correct detection trials in the first and fourth quartiles of the RT distribution. On the average, correct detections required 235.3 ms more time on low-performance trials (quartile 4) than on high-performance trials (quartile 1).

TABLE 2. Mean Reaction Times to Correctly Detected Stimuli Sorted by RT Quartile.

Quartile	Mean	SD
1	335.94	49.01
4	571.28	125.71

GRAND AVERAGE ERPS

Figure 2 contains high- and low-performance ERPs from Fz, Cz, and Pz, averaged across trials and subjects. The reference electrode for the traces in Fig. 2 was at the nose. Voltages are expressed relative to machine zero (amplifier DC).

The initial deflection visible in the responses is N1, which is the negative peak near 100-ms poststimulus. As is usually the case, N1 is

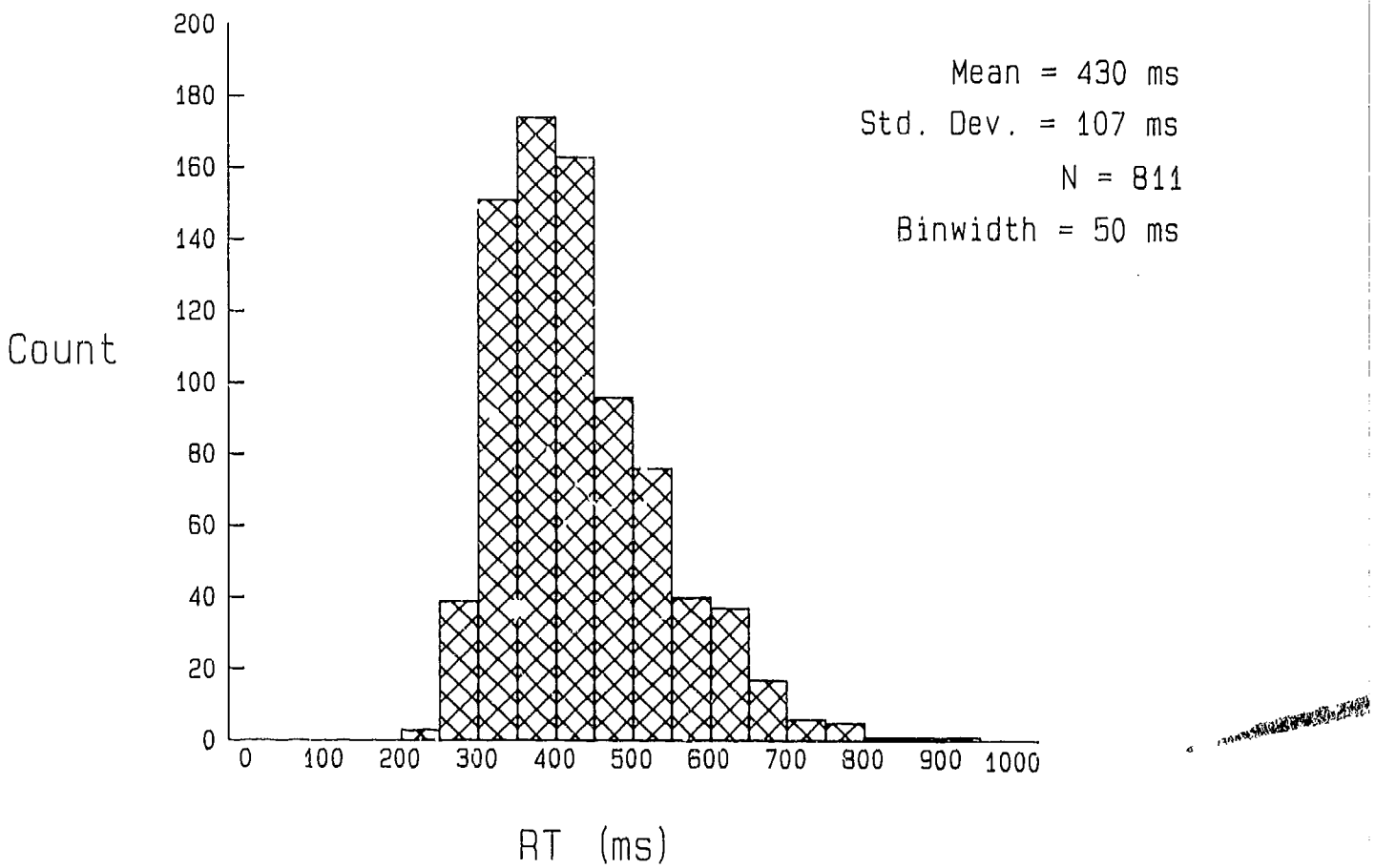

Figure 1. The overall distribution of reaction times (RTs) corresponding to correctly detected stimuli.

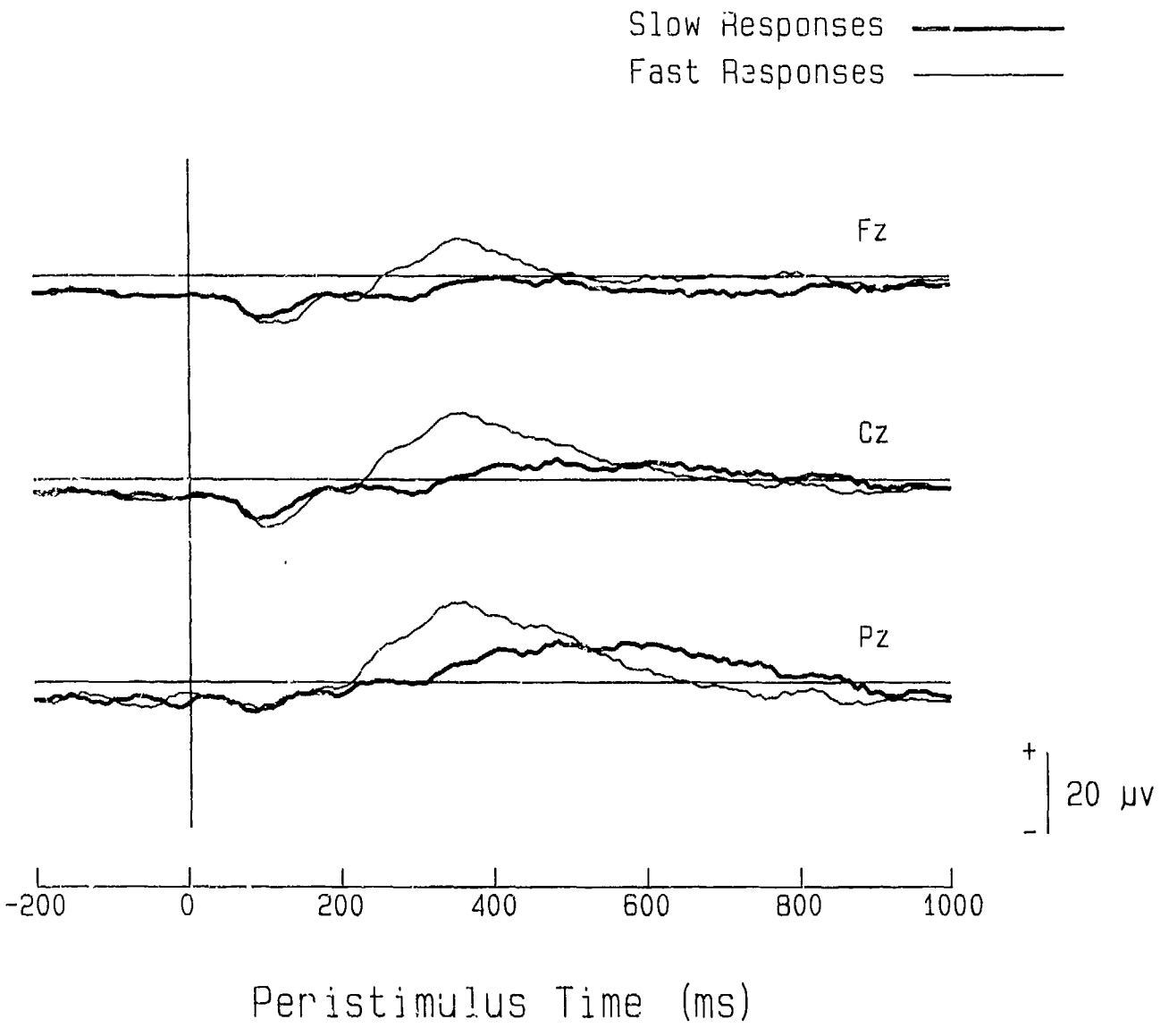

Figure 2. Grand average ERPs from Fz, Cz, and Pz corresponding to the first and fourth reaction time quartiles. The vertical line drawn through the traces indicates stimulus onset. The responses from different recording sites are offset, vertically, for clarity. The horizontal line drawn through each pair of traces indicates machine-zero voltage (amplifier-DC).

largest centrally and frontally and is somewhat smaller parietally. What appears to be a small processing negativity can be seen overlapping the trailing edge of the high-performance N1 responses (giving the high-performance N1 the appearance of being extended in duration).

Positive-going P2 responses can be seen peaking near 170-ms poststimulus at Fz and Cz in the quartile 1 responses. These are followed by negative-going (and somewhat poorly defined) N2 peaks at about 220-ms poststimulus.

The LPC is evident in the quartile-1 responses as a large, positive-going wave peaking at about 360-ms poststimulus (see Table 3). The peak amplitude of the quartile-1 LPC is largest at the parietal electrode and diminishes anteriorly, the usual LPC topography. The quartile-4 LPC is a broad positivity, except at the frontal electrode, where the averages are never positive during the recording interval. The quartile-4 LPC topography is qualitatively similar to that of the quartile-1 LPC. The quartile-4 LPC maximum occurs at an average latency of about 488 ms poststimulus.

TABLE 3. Mean LPC Amplitudes and Latencies--Grand Average Response Measurements.

Quartile	Fz		Cz		Pz	
	Ampl.	Lat.	Ampl.	Lat.	Ampl.	Lat.
1	9.2	356	16.6	364	19.9	364
4	9.2	484	5.2	488	10.2	492
Difference	9.2	-128	11.4	-124	9.7	-128

LPC PEAK MEASUREMENTS

Within-subject Signal Averages at Fz, Cz, and Pz

As in the grand averages, the within-subject LPC peak amplitudes were largest at the parietal electrode and diminished anteriorly (see Table 4). The test of the effect of electrode position confirms this observation ($F = 43.98$, G-G $df = 1.83$, $p < .00005$).

The LPC peak amplitudes in the within-subject averages were much larger on high performance trials than on low performance trials (averaging 19.1 versus 10.1 uV, overall; $F = 28.38$, $df = 1,9$, $p = .0005$). The average change in response amplitude between RT quartiles was somewhat larger at Cz than at Pz and rather smaller at Fz. This pattern, however, as measured by the electrode by quartile interaction, was nonsignificant.

The LPC latencies in the within-subject averages were shorter on high-performance trials by an overall average of 103.1 ms. The effect of performance level on LPC latency was significant ($F = 38.49$, $df = 1,9$, $p = .0002$). The magnitude of the latency shift between the two performance

levels varied substantially between the three electrodes (F = 7.9, G-G df = 1.44,12.92, p = .009), amounting to 155.6 ms at Pz, 118.8 ms at CZ, and 44.8 ms at Fz.

TABLE 4. Mean LPC Amplitudes and Latencies--Within-subject Signal Averages.

Quartile	Fz		Cz		Pz	
	Ampl.	Lat.	Ampl.	Lat.	Ampl.	Lat.
1	12.3	372	21.0	376	24.1	356.8
4	4.4	407.6	10.8	494.8	15.3	512.4
Difference	7.9	-34.8	10.2	-118.8	8.8	-155.6

Regression Analysis of Reaction Times and LPC Latencies

A preliminary analysis of the single-trial LPC latencies was carried out to compare the results of this experiment with those of several existing studies that have examined correlations between RTs and single-trial LPC latencies. Figure 3 shows the overall distributions of LPC latencies at the three recording sites. The LPC latency distributions are quite similar to the distribution of RTs (replotted in Fig. 3, for comparison). Goodin and Aminoff (18) reported auditory LPC latency distributions that were more symmetrical than ours (and more symmetrical than their RT distributions).

Figure 4 contains a scatterplot of RT versus LPC latency at the Pz electrode for the individual trials from all 10 subjects and all 4 quartiles of the RT distribution. Examining the plot closely, one may be able to see two clusters of data points. One cluster roughly follows the unit slope line (which has been drawn in the figure). The other forms a broad and somewhat faint horizontal stripe with a center of gravity near 400 ms.

The second cluster of points appears to represent failures of the peak-picking algorithm. Their distribution is consistent with the existence of a subset of trials on which low signal-to-noise ratios rendered the peak-detecting algorithm's output essentially noise. Indeed, if the peak-finding algorithm were applied to noise, the marginal distribution of latencies would be uniform on the horizontal axis, and the RT marginals would be distributed as in Fig. 1. The resulting scatterplot would be a blurred, horizontal stripe, densest near the mode of the RT distribution. The second cluster of points just mentioned appears to fit this description. Evidence of this phenomenon also can be seen in the data of Pfefferbaum et al. (19). Note that the axes in the figures of Pfefferbaum et al. are reversed relative to those of Fig. 4 of this report. Those investigators cast latency as the criterion variable, rather than RT, hence the stripe in their figure runs vertically.

The line drawn through the points of Fig. 4 with a slope of 0.35 is the least-squares solution obtained from the regression of RT on LPC latency.

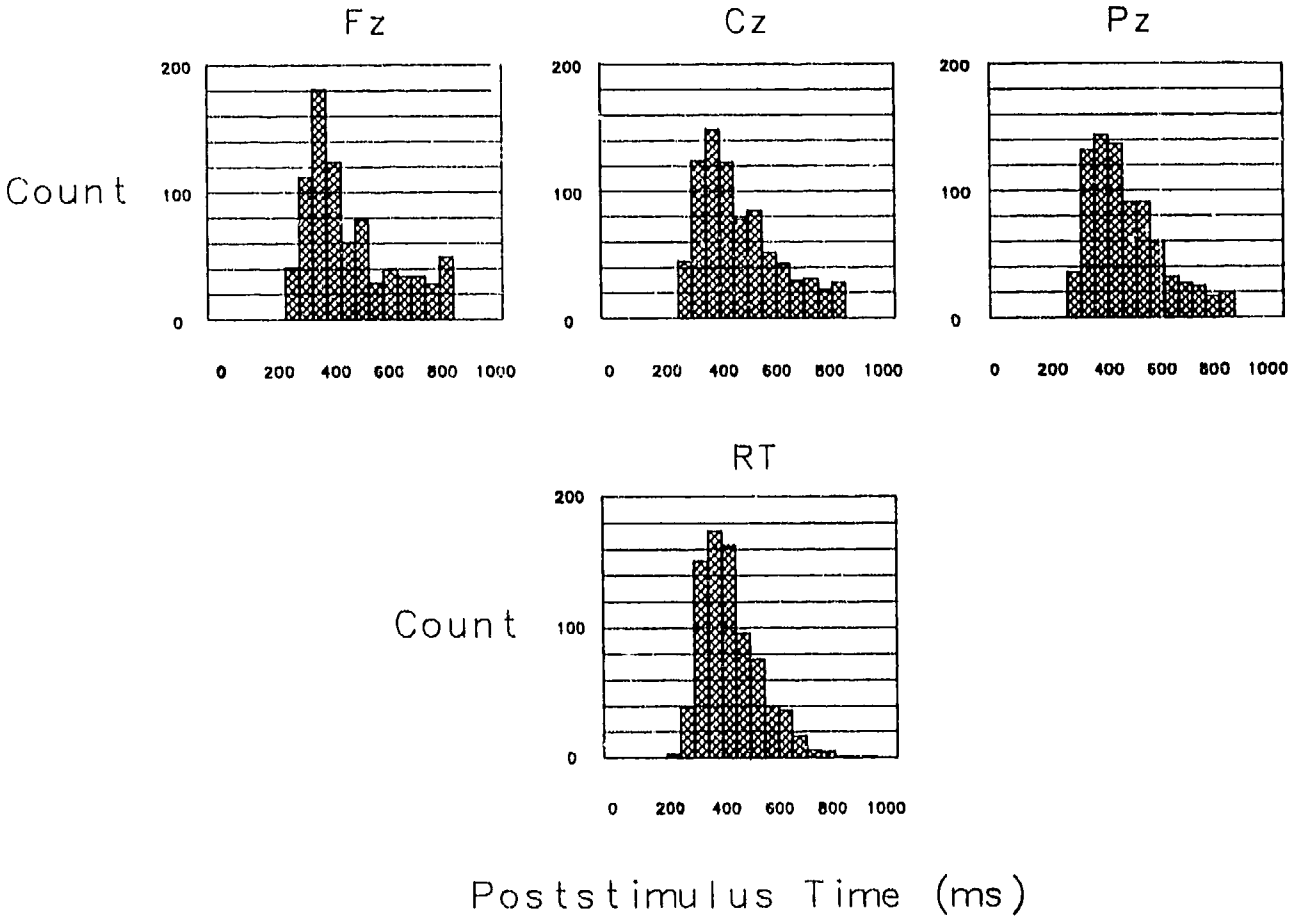

Figure 3. Frequency distributions of the single-trial LPC latencies at Fz, Cz, and Pz recording sites. The reference electrode was at the nose. The overall RT distribution is redrawn in the figure, for comparison.

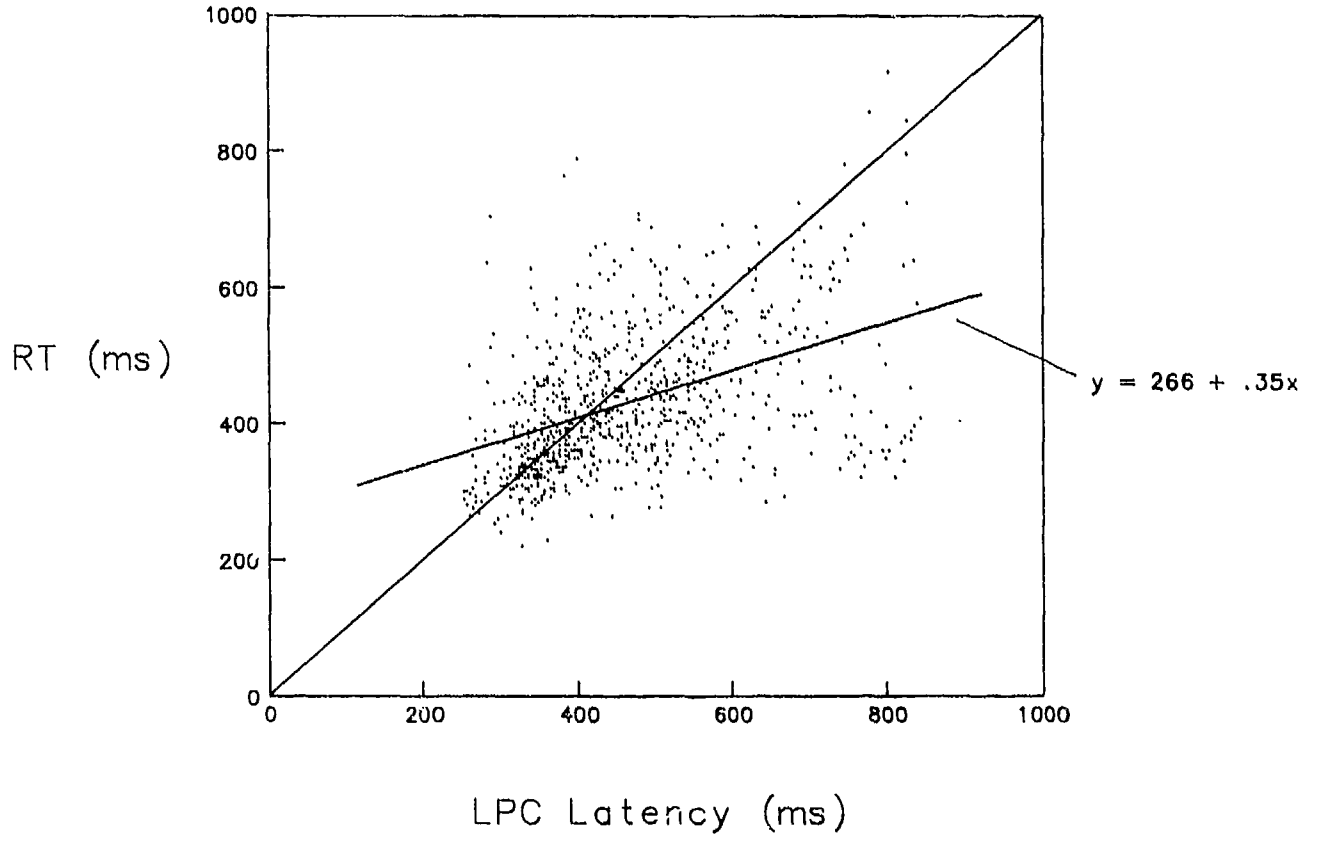

Figure 4. Single-trial reaction times plotted versus LPC latency at the Pz electrode (nose reference). The line with -0.35 slope is from the least squares regression of reaction time on latency. The unit slope line is drawn in the figure for comparison.

Examining the figure in light of the foregoing suggests that the slope of the least-squares line is a compromise solution, rather shallower than it might have been had it been possible to eliminate the presumably badly estimated latencies discussed previously.

The overall bivariate correlations between RT and LPC latency at Fz, Cz, and Pz were .22, .36, and .44, respectively. The nominal Type I error rate in each case is $p < .0005$, which increases to $p < .0015$, after Bonferroni adjustment for three, simultaneous tests (the p values are also nominal for reasons that will be discussed presently).

The values of these correlations are in reasonable agreement with the values reported by Pfefferbaum et al. (19), who obtained a maximum overall correlation at Pz of .30 in speeded visual discrimination. Their latency measurements were derived using a cross-correlational technique (by a Woody (20) style adaptive filtering technique). In a study of the effects of nitrous oxide, Fowler et al. (21) reported somewhat higher correlations between RT and P3 at Pz, ranging from .51 to .66. Those authors also used a Woody-style adaptive filter to obtain their latency estimates.

The full-rank multiple regression of RT on the three latencies, and their squared values (included as a check for nonlinearity), yielded a correlation of .48 (nominal $p < .00005$). The correlation obtained is only slightly larger than the bivariate correlation between RT and latency at the Pz electrode. This result suggests that LPC latency measurements might have little predictive value beyond that afforded by the Pz electrode alone.

As noted previously, the probability levels associated with the overall correlation coefficients just discussed may not be accurate. That analysis treated 10 sets of approximately 80 x-y pairs (obtained from 10 different subjects) as if they contained approximately 800 independent observations, which probably was not the case. To verify that the correlations just discussed were indeed real, the sample of RTs and latencies was split into two randomly selected subsamples of equal sizes (sampling was without replacement). One subsample was used to calculate an equation for predicting RT from LPC latency. This equation was then applied to the second sample to predict RTs from latencies.

Figure 5 shows the results obtained from 500 random cross validations carried out in this way. The average value of the RMS prediction errors in the new samples was 96.5 ms; the obtained value never exceeded the overall standard deviation of the RTs. (The overall standard deviation of the RTs is the best sample estimate of the corresponding population parameter; that, in turn, is the standard error of the estimate one would expect to obtain from the regression if the LPC latencies provided no information about RT.) Inserting the 2.5- and 97.5-percentile values of the Monte Carlo standard errors into the formula for the correlation coefficient, and assuming a value of 107 ms for the standard deviation of RT, yields a Monte Carlo 95% confidence interval for the correlation coefficient in <u>new</u> samples that ranges from a lower limit of .27 to an upper limit of .54. This suggests that the latency-RT correlation is real, replicable in new samples of data, and similar in magnitude to correlations reported by others (19,21).

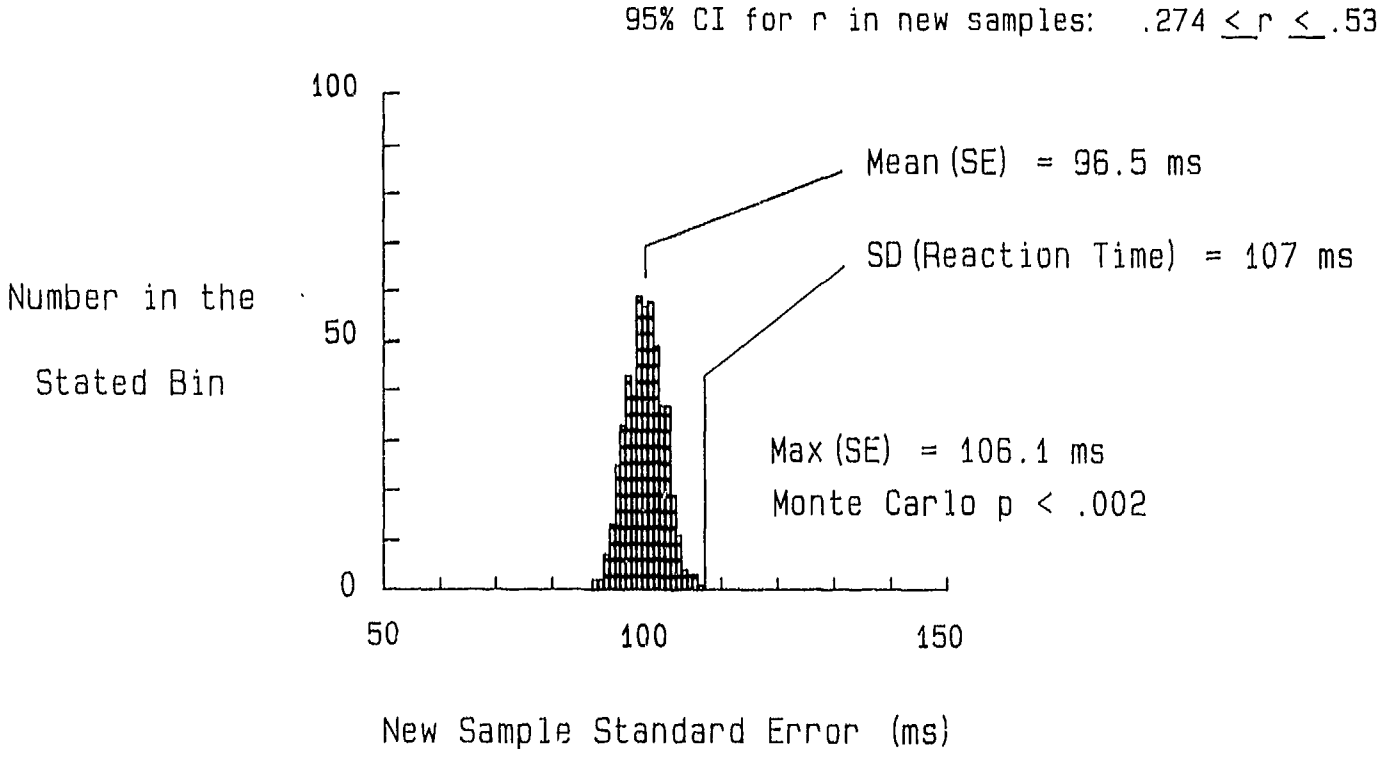

Figure 5. The distribution of standard errors of the estimate from 500, random, split-half cross validations of the regression of LPC latency on RT. The "new sample" standard errors shown are standard errors of the estimate that resulted when the regression line calculated using one half of the data set was used to predict RTs in the other half of the data.

Figure 6 contains scatterplots of RT against LPC latency at Pz broken out by subject. One can see that the strength of the relationship between RT and latency varies substantially among individuals. The associated correlations for the P3 electrode range from .03 to .71 (mean = .38, median = .39). These values are lower than the individual subject correlations from the auditory oddball experiment reported by Ritter et al. (22). Correlations ranged from .50 to .79 (mean = .66, median = .68) when the data were handscored (22).

In 3 of the 10 subjects, the correlation between RT and LPC latency was greatest at either Fz or Cz rather than Pz. The differences among the correlations for the different sites appear to have been uniformly negligible: The full-rank multiple regression of RT on latency at the three recording sites yielded an average multiple correlation of .43. Given the small size of the partial correlation and the increased opportunity to capitalize on chance afforded by multiple regression, the improvement in predicting RT from LPC latencies attributable to increasing the number of electrodes from one to three appears trivial, at least in this sample of subjects.

Analysis of Single Trial LPC Measurements Sorted by RT Quartile

Sorting the trials by RT quartile yielded single-trial LPC amplitudes that were largest at the parietal electrode and diminished anteriorly, thus mirroring the grand and within-subject averages (Table 5; main effect F = 44.55, G-G df = 1.45,13.04, p = .0003). The amplitudes of the LPCs measured in single trials are uniformly larger than those measured in the within-subject averages (which, in turn, are larger than those measured in the grand averages). At least two reasons are possible. One is the fact that when single-trial latencies are variable, signal averaging will smear the peaks in the individual trials, yielding an average peak whose amplitude is less than the mean of the single-trial peaks' amplitudes.

The second reason is that peak-finding algorithms that are basically waveform maximum detectors (such as the one used here) are biased in the presence of noise. The bias is due to the fact that the noise levels of single trials are higher than those of averages. When a response contains signal and noise, the maximum of the signal-plus-noise waveform will tend to occur at a point where the noise voltage is large and of the same polarity as the signal. This tends to bias peak amplitude measurements toward more extreme values. The bias toward large amplitudes decreases as noise decreases. Hence, it is smaller in signal averages.

The peak amplitudes of responses associated with rapid detections were larger than those of slow detections (F = 24.7, df = 1,9, p = .0001), as occurred with the within-subject and grand averages. The performance-related amplitude differences, however, were smaller in the single trial measurements than in the averages. The mean effect of performance level on LPC amplitude in the single trials was largest at Fz, if anywhere. The electrode by RT-quartile interaction was nonsignificant here, as it was for the signal averages (see Table 5).

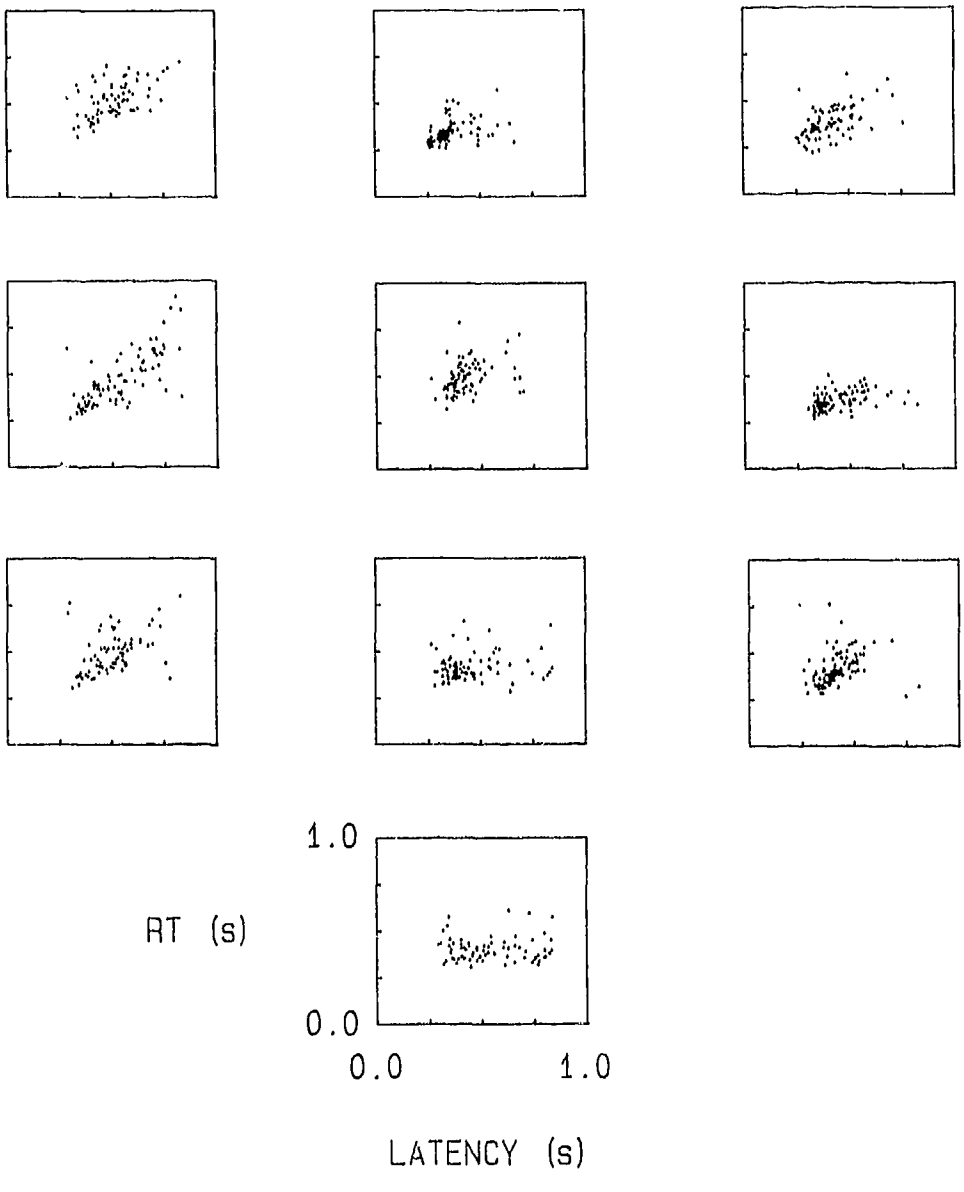

Figure 6. Single-trial reaction times plotted versus LPC latency at the Pz electrode for each subject (nose reference).

TABLE 5. Mean LPC Amplitudes and Latencies--Single-trial Measurements.

Quartile	Fz		Cz		Pz	
	Ampl.	Lat.	Ampl.	Lat.	Ampl.	Lat.
1	20.61	455.0	29.41	416.5	32.79	408.3
4	14.89	492.5	24.37	503.6	28.94	493.7
Difference	5.72	-37.5	5.04	-87.1	3.85	-85.4

In the single trials, LPC latencies were again shorter on high-performance trials than on low-performance trials (Table 5): The overall difference averaged 70.0 ms across electrodes ($F = 19.81$, $df = 1,9$, $p = .0018$). The topography of the single-trial latency shifts differed from that observed in either type of signal average, however, yielding a significant electrode by RT-quartile interaction ($F = 6.74$, G-G $df = 1.71,15.43$, $p = .0099$). The performance-related shifts in LPC latency were about equal at Cz and Pz in the single trials (Table 5). As in the within-subject averages, the latency shift at Fz was comparatively small.

LPC MEASUREMENTS: AVERAGE AMPLITUDE FROM 300 TO 500 MS POSTSTIMULUS

The time-averaged response amplitudes recorded at Fz, Cz, and Pz during an interval from 300 to 500 ms poststimulus are shown in Table 6. Again, LPC amplitudes were largest at Pz and progressively smaller at more anterior recording sites. This pattern differs reliably from a flat topography ($F = 48.56$, G-G $df = 2.0,17.99$, $p < .00005$). The amplitude difference among performance levels, which averages 9.03 uV, is largest at Cz, smaller at Pz, and considerably smaller at Fz (Table 6). The overall difference is significantly nonzero ($F = 14.85$, $df = 1,9$, $p < .005$). The difference in magnitude of the performance effect across recording sites is evidently reliable, yielding a significant electrode by quartile interaction ($F = 12.29$, $df = 2,8$, $p < .004$).

TABLE 6. Average LPC Amplitude Measured from 300 to 500 ms Poststimulus.

Quartile	Fz	Cz	Pz
1	4.98	12.46	15.75
4	-1.70	1.84	5.94
Difference	6.68	10.34	9.81

BIPOLAR RECORDING MONTAGE: RMS-S ANALYSIS

Figure 7 shows grand averages of the responses obtained from the bipolar derivations. All of the responses in the figure are drawn so that positivity at the parietal electrode is downward.

Notable features of these responses are the broad, negative-going slow waves of perhaps 600-700 ms duration in the anterior-posterior recordings (the first four sets of responses in Fig. 7). The slow waves reach their maximum absolute amplitudes in the 350-ms latency range in the high-performance quartile averages.

All of the bipolar responses contain minor waveforms that appear to be N1 responses. These minor waves span the interval from roughly 60 ms to near 200 ms poststimulus. They reverse polarity in the occipital-parietal traces, which suggests a generator or generators anterior to the P3 electrode.

The high-performance responses tend to be parietally positive (relative to the low-performance responses) in the 350-ms latency range. Unlike the 60-200 ms interval, the 350-ms region contains no evidence of a polarity reversal. This suggests a generator with a parietal voltage maximum, which is consistent with the idea that this portion of the waveform reflects LPC activity.

The low-performance responses reach their absolute maxima nearer 600 ms poststimulus and appear to contain a frontally distributed negativity rather than a parietal positivity. The low-performance waves tend to become relatively positive in the occipital-parietal derivations in the 700-1000 ms region.

The overall pattern of slow wave activity appears to be reliable. An analysis of variance calculated using RMS-s statistics derived using a pair of intervals spanning 0-500 ms and 500-1,000 ms poststimulus, respectively, yielded a significant windows by quartile interaction ($F = 19.83$, $df = 1,9$, $p = .0016$). These time windows were chosen after inspecting the data, however, a fact that clouds their significance levels.

Table 7 shows the results of a set of analyses of variance carried out on RMS-s statistics calculated using 50-ms windows (after the method of Trejo et al. (1)). Each ANOVA was a 2 by 6, quartiles by electrodes, within-subjects design. Each F statistic in the table corresponds to the main effect of quartiles during the indicated time window. Each p value is the comparisonwise Type I error rate for its associated F ratio. The largest effect of performance level occurs during the LPC time course in the intervals spanning 300-448 ms poststimulus. The effect is maximum in the 300-348 ms interval.

FZ, CZ, PZ MONTAGE: RMS-S ANALYSIS

To compare the bipolar and Fz, Cz, Pz electrode arrays, we also calculated RMS-s statistics for the latter using the same group of approximately 50-ms time windows used in the analysis of the bipolar responses. Figure 8 shows the grand average signal-to-noise ratio vectors

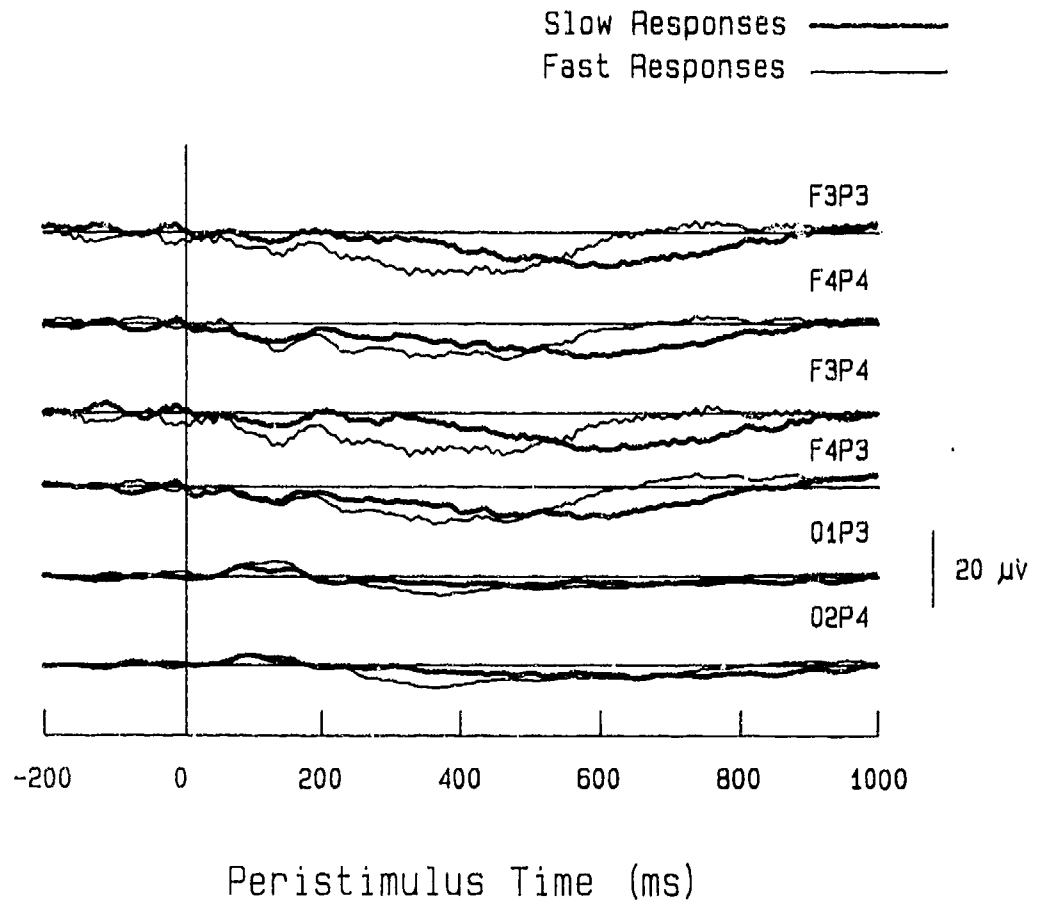

Figure 7. Grand average ERPs obtained with the bipolar montage, sorted by reaction-time quartile. The vertical line drawn through the traces indicates stimulus onset. The responses corresponding to different electrode pairs are offset vertically for clarity. The horizontal line drawn through each pair of traces indicates zero potential difference between the two electrodes corresponding to that derivation. In each pair of traces, an upward deflection indicates positivity relative to the reference electrode.

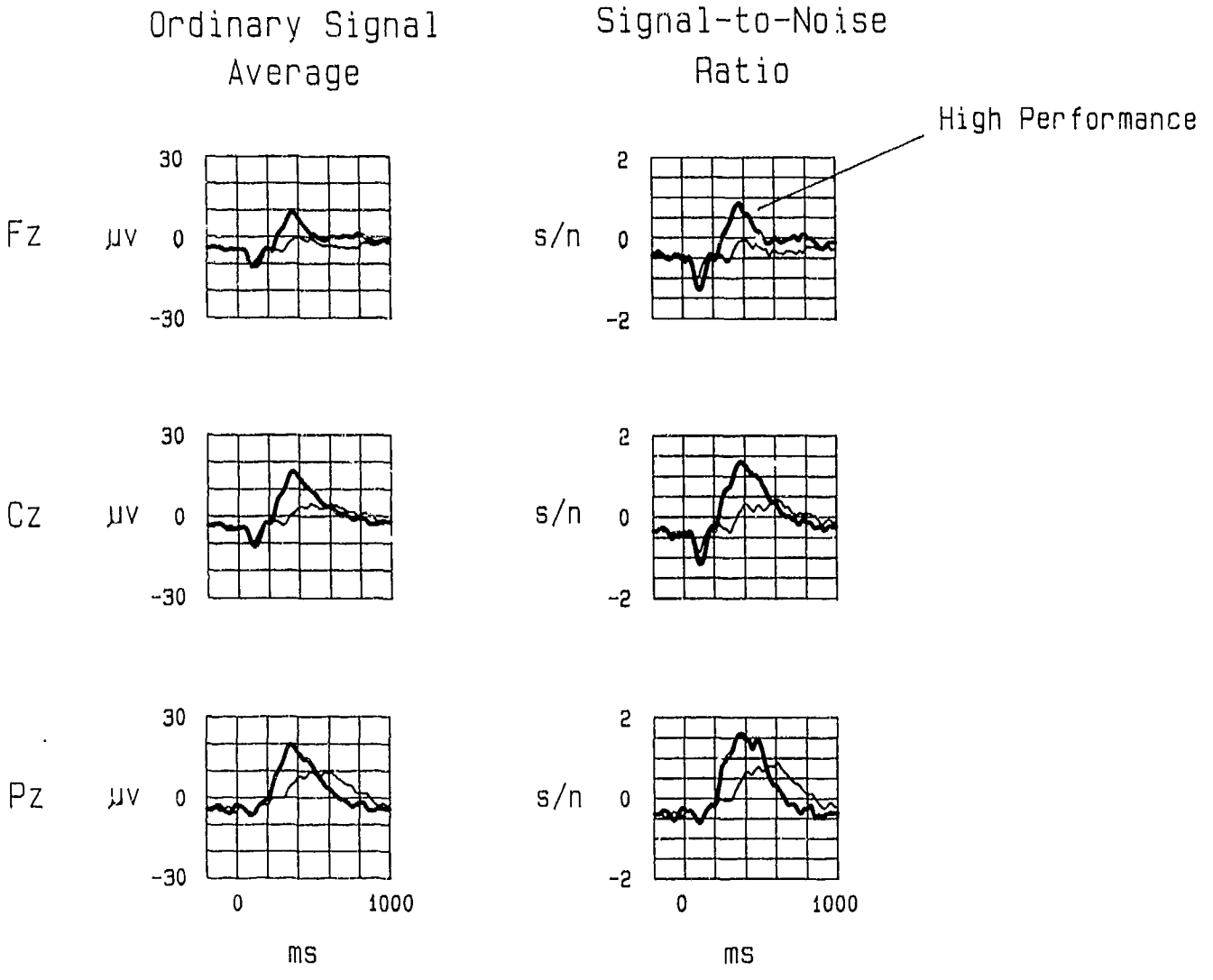

Figure 8. A comparison of grand signal averages and grand average signal-to-noise ratios (S/Ns), after sorting trials by reaction-time quartile. High-performance (quartile 1) averages are drawn with heavy lines. The reference electrode was at the nose.

from Fz, Cz, and Pz on which the RMS-s calculations were performed. These vectors do not differ markedly from the grand averaged Fz, Cz, and Pz ERPs.

TABLE 7. F Ratios and Comparisonwise Type I Error Rates (p) for the Effect of Performance Level on Bipolar Response RMS-s Statistics.

Analysis time window (ms poststimulus)	F	p
48-100	1.16	.3103
100-148	2.23	.1699
148-200	2.45	.1522
200-248	0.26	.6253
248-300	2.72	.1333
300-348	14.38	.0043
348-400	8.42	.0176
400-448	7.31	.0243
448-500	1.17	.3072

A more detailed comparison of S/N waves and signal averages is shown in Fig. 9. Each point in the scatterplots of Fig. 9 is the grand mean signal-to-noise ratio at a given poststimulus time, t, plotted against the grand mean voltage at time t. The six panels in the figure correspond to the three recording sites and two performance levels. Examining Fig. 9, one can see that the points tend to cluster along the 45-deg lines (which, due to the difference in the scales of the two axes, correspond to slopes of 0.1).

Because the mean voltage at time t is the estimated signal level at time t, a linear relation between voltage and S/N indicates that the average noise level is constant across time. The linear appearance of the scatterplots suggests that S/N is fairly constant across time, except at the higher voltages in the high performance ERPs where the points tend to form loops. The lower segments of the loops correspond to transient reductions in S/N during the early portion of the LPC (see the responses in Fig. 8).

The F ratios for the main effects of performance levels on the RMS-s statistics are shown in Table 8, along with their associated, comparisonwise error rates. A difference between the results for the bipolar recordings and those for the Fz, Cz, Pz-montage is that the latter montage yields larger performance effects during the first 250-ms poststimulus (during the N1, P2, and N2 waves). As in the bipolar responses, however, the largest effect of performance occurs during the 300-348 ms time window (during the LPC time course). The magnitude of this effect is much larger in the Fz, Cz, Pz recordings than in the bipolar recordings.

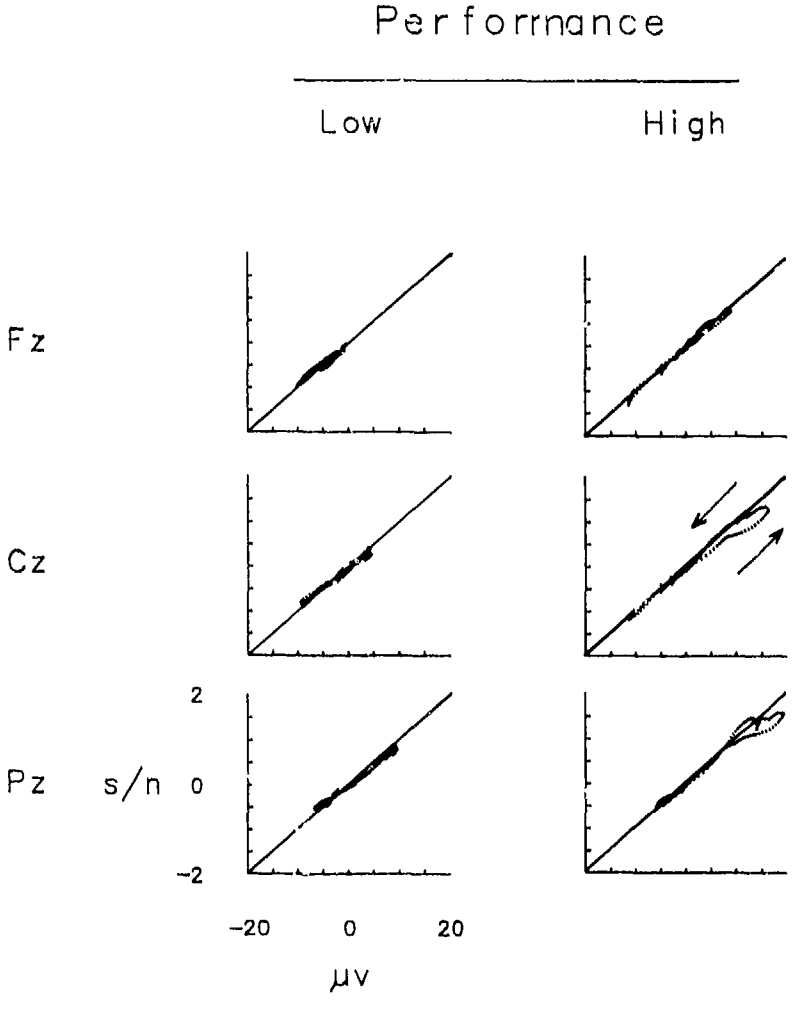

Figure 9. Mean signal-to-noise ratio plotted against mean response voltage. The ordinate of each plot is mean ERP signal-to-noise ratio, averaged across trials and subjects. The abscissa is mean ERP amplitude, similarly averaged. Each point corresponds to the signal-to-noise ratio at one peristimulus time. The plots in the left-hand column are for trials in RT quartile 4 (low performance); the plots in the right-hand column are for trials in RT quartile 1 (high performance). The reference electrode was at the nose.

TABLE 8. F Ratios and Comparisonwise Type I Error Rates (p) for the Effect of Performance Level on RMS-s Statistics from Fz, Cz, and Pz.

Analysis time window (ms poststimulus)	F	p
48-100	3.30	.1026
100-148	3.74	.0295
148-200	7.69	.0216
200-248	7.59	.0223
248-300	1.97	.1940
300-348	41.56	.0001
348-400	8.27	.0183
400-448	6.20	.0344
448-500	3.60	.0904

COMPARISON OF NOSE AND VERTEX REFERENCES

We compared the abilities of nose- and vertex-referred recordings to distinguish between performance levels. The recording sites used in this analysis comprised a symmetrical, lateral array consisting of F3, F4, T3, T4, P3, P4, O1, and O2. Note that this is the same set of recording sites used in the bipolar analysis. The comparisons were performed using the RMS-s statistic calculated on the nine 50-ms time windows discussed earlier.

Grand averaged responses obtained using the two references are shown in Fig. 10. Responses from vertex, referred to nose, have been included in the left-hand column of traces for illustration.

Comparing the two sets of traces, one first notes that N1 is inverted in the vertex-referred data, a result attributable to the fact that the N1 is negative across most of the scalp surface and reaches its maximum at vertex (see the nose-referred traces). The performance related effects in the ERPs, which are large at posterior recording sites in the nose referred traces, are reduced, inverted, and displaced frontally and temporally in the vertex referred data.

Results from the statistical analysis of the RMS-s measurements are shown in Table 9. Examining Table 9 confirms that using the vertex reference yields LPC performance effects that are much reduced, relative to those in the nose-referred data. Interestingly, the effect of performance on N1 exceeds the effect of performance on the LPC in the vertex-referred data (compare the 148-200 and 300-348 ms intervals in Table 9; compare also the traces of Fig. 10).

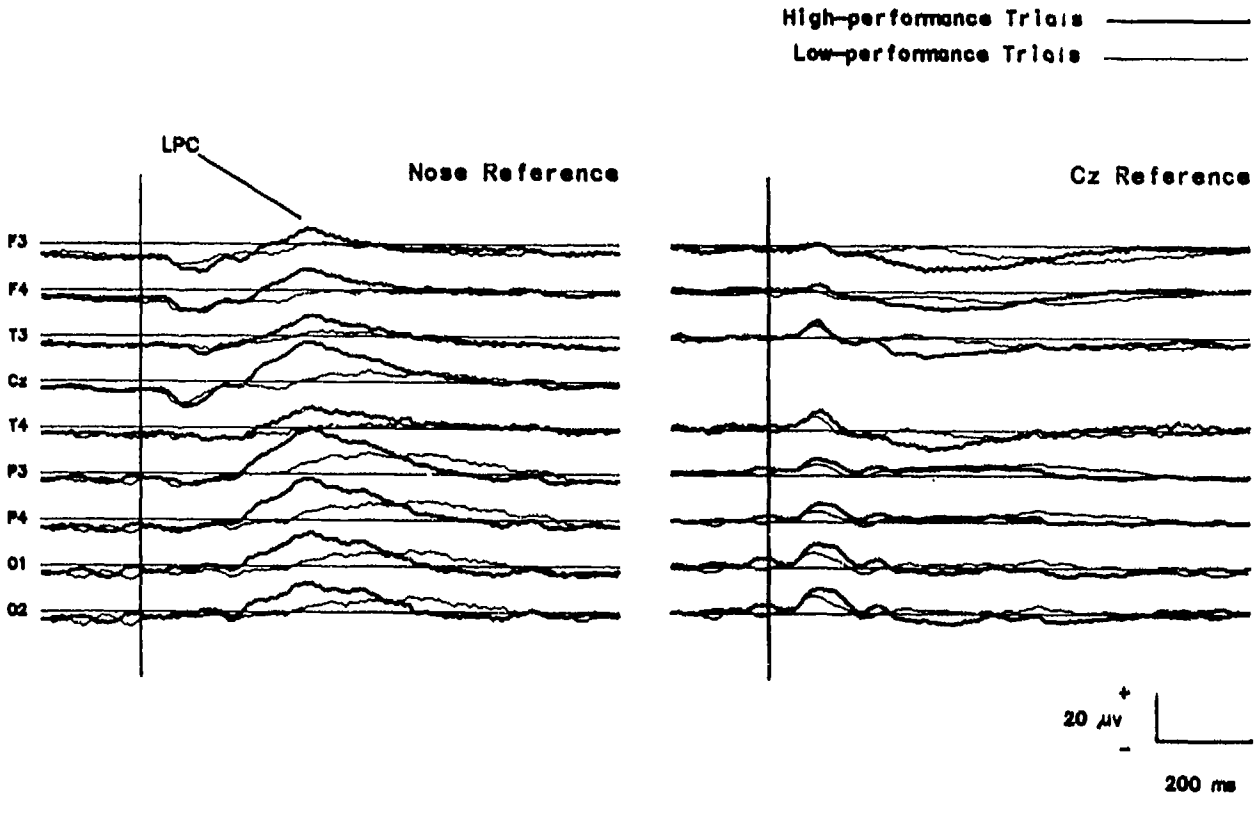

Figure 10. A comparison of nose- and vertex-referred waveforms. The averaged ERPs shown correspond to trials in the first and fourth RT quartiles and are averaged across trials and subjects. The ERPs in the left-hand column are referred to nose; the vertex potential (re nose) is included in this column for comparison. The ERPs in the right-hand column are referred to vertex. Horizontal lines indicate amplifier DC voltage in the nose-referred traces and voltage relative to vertex in the Cz-referred traces.

TABLE 9. F Ratios and Comparisonwise Type I Error Rates (p) for the Effect of Performance Level on RMS-s Statistics Derived using the Lateral Electrode Array with either Vertex or Nose Reference.

Analysis time window (ms poststimulus)	Reference			
	Vertex		Nose	
48-100	2.52	.15	1.15	.31
100-148	7.43	.02	4.98	.05
148-200	5.78	.04	6.69	.03
200-248	1.06	.33	1.06	.33
248-300	1.04	.33	1.51	.33
300-348	4.77	.06	45.15	.00
348-400	2.91	.09	14.75	.01
400-448	2.73	.04	8.31	.02
448-500	0.04	.85	2.00	.05

As noted elsewhere, the effect of performance on the LPC is usually large at vertex. For that reason, and because the LPC is broadly distributed across the scalp, the process of subtracting the vertex potential from each of the other electrodes (which is what referring them to vertex accomplishes) can be expected to attenuate correlations between the physiological and behavioral data.

Of note, none of the electrodes by response-quartile interactions was significant in the analyses of variance performed on the vertex-referred RMS-s values. Hence, the comparatively small F ratios yielded by the Cz-referred data cannot be attributed to a failure of additivity.

DISCUSSION

The results of this study replicate and extend previous observations (13-15). The most pertinent result is that short-term changes in performance are accompanied by short-term changes in the properties of ERPs--in particular, by changes in the properties of the LPC.

One clear result of the present study was that LPC amplitudes discriminated better than LPC latencies among performance levels. Peak amplitudes measured in signal averages discriminated better than peak latencies measured in signal averages, and amplitudes measured in single trials discriminated better than latencies measured in single trials. Time-averaged LPC amplitudes calculated from 300-500 ms poststimulus performed less well; however, RMS-s measurements (which are, in effect, weighted time averages) taken 300-348 ms poststimulus yielded the best discriminations of all.

Although amplitudes discriminated better among performance levels, the association between LPC latencies and RTs was very reliable, nonetheless.

Thus, nothing in our data contradicts the idea that LPC latencies index the latencies of perceptual decisions (23,24).

Judging from the RMS-s analysis, N1 amplitude discriminated performance levels less well than LPC amplitude (except in the vertex-referred data). An effect of performance on N1 was present; however, the effect was much smaller than the effect on P3.

The good performance of the RMS-s statistic warrants discussion. The RMS-s transforms calculated from approximately 300 to 350 ms poststimulus (which is during the LPC time course) outperformed signal average peak amplitudes, signal average peak latencies, single-trial peak amplitudes, single-trial peak latencies, and 300-500 ms average amplitudes.

One can argue that the large number of intervals during which the RMS-s was calculated afforded the opportunity for one interval to yield significant results by chance. Given the empirical association between LPC amplitude and performance (reviewed in 25), however, it would have been very surprising if response magnitudes in this interval had not proven sensitive to variation in performance level.

A second question with regard to the RMS-s statistic concerns whether its advantage over single-trial measurement was due to the effects of signal averaging. (Recall that the RMS-s was calculated on five-trial signal averages.) That the RMS-s statistic also outperformed peak measurements taken on 20-trial signal averages suggests this was not the case.

One also might argue that the advantage of the RMS-s statistic over peak measurements observed in this experiment might be method-dependent rather than general. The algorithm used to obtain peak voltage measurements in this study was simple, and more elaborate procedures might have yielded better results. A comparison of the F statistics corresponding to the main effect of performance indicates that the power advantage of the RMS-s statistic (in the 300-348 ms poststimulus interval) is equivalent to a 68% reduction in error variance, relative to that of the single-trial amplitude measurements. Single-trial measurements might be improved correspondingly, but such an improvement would be substantial. Of note, the correlation between LPC latency and reaction time obtained in the present study appears to be well within the range of values obtained by cross correlational adaptive filtering (Woody filtering (20); see references 19 and 21). This suggests the improvement afforded by that technique may not always be large.

A final point with regard to the RMS-s transform concerns the degree to which the results it produces differ from other time-integrated voltage measurements. No extensive comparisons were made in this study, but Trejo et al. (1) compared the RMS-s and ordinary RMS amplitude measures and reported little difference between the two. This is consistent with the comparison between signal-average and signal-to-noise ratio vectors presented here, the results of which suggest that the differences between the two are minor. Evidently, RMS and RMS-s transforms produce similar results because the signal average vectors used to calculate RMS amplitudes and the signal-to-noise ratio vectors used to calculate RMS-s magnitudes tend to be proportional to one another across most of their ranges (see Fig. 9).

Theoretical considerations suggest that a simple voltage average across time should yield amplitude estimates with smaller measurement errors than estimates from RMS-based techniques: Zero-mean noise superimposed on a signal tends to cancel in averages calculated on voltages, whereas it sums with the signal in amplitude measures calculated on squared voltages (such as RMS voltages). This advantage is counterbalanced, however, by the fact that the average voltage can be zero if the signal's polarity is not constant. Thus, when constant signal polarity can be assumed, the simple average of a time series may yield better amplitude estimates in the sense that errors will be smaller. When constant polarity cannot be assumed, an RMS transform (or some other rectifying transform) may be a better choice.

The comparatively poor performance of the 300-500 ms time average seems to be due to the length of the averaging interval. An examination of the F ratios corresponding to the windowed RMS-s statistics for the array of midline electrodes (Table 9) indicates that the correlation between amplitude and performance level decreases rapidly after 300-350 ms poststimulus and approaches zero by 450-500 ms poststimulus. Indeed, the average of the t statistics corresponding to RMS-s measurements in the intervals spanning 300-500 ms poststimulus is 3.43, which is near the value of 3.86 obtained from the 300-500 ms simple average. (These t values are the square roots of the F ratios corresponding to the performance main effects; they might be interpreted as mean distances between the ERP measurements at high and low performance, expressed in units proportional to the intersubject variability.)

Judging by their relative sensitivities to the effects of performance on LPC amplitude during the 300-350 ms interval, two of the electrode montages examined performed in superior, and nearly equal, fashion. These were the lateral and midline arrays used with the nose reference. Of these, the midline array performed at about 96% of the level of the larger, lateral array in terms of the distance measure defined previously. The bipolar montage was a distant third, at 56% of the performance of the best array. The lateral montage used with the vertex reference was last, at 32% of the performance of the best montage.

A consideration of the foregoing suggests that the Fz, Cz, Pz montage, referred to nose, was arguably the most efficient recording array, in that it yielded by far the most information per electrode. Indeed, its absolute performance was only marginally inferior to the larger lateral array.

Theory suggests that P3 is a reasonable choice as a general index of performance; it does not appear to be specific to any sensory modality, and it varies systematically with performance in detection, attention, and other cognitive processes (see reference 25). Hence, the P3 generators perform either a very wide array of functions or a few very general ones. In either case, these response properties make P3 a useful index of the aggregate performance of a number of cognitive mechanisms.

The aggregate nature of P3 probably also limits its practical value as a performance index. The P3 responses recorded from the scalp surface appear to be comprised of perhaps seven or more overlapping positive waves. These presumed underlying waves include P3a, P3b (P3 "proper"), P3e, Px, the frontal P3, the positive Slow Wave(s) of parietal scalp, and perhaps

others (10-12). The components of the P3 family do not necessarily behave similarly. For example, P3b (which may be the P3 usually studied) and the tonic Slow Wave bear opposite relationships to behavior: P3b tends to be large when a task is easy and decisions are confident and small when a task is difficult and decisions are unsure. Slow Wave, on the other hand, tends to behave oppositely (10-12). Due to this confounding, an increase in task difficulty might increase a global measure of P3 amplitude, decrease it, or leave it unchanged, depending on the relative contributions of the two underlying responses.

Presently, we have no valid way to distinguish between simultaneous responses without depending on knowledge of the locations of their sources (6). Existing data suggest that motor responses can precede P3 responses (18), which in turn suggests that at least some of the decisions that inform motor responding can be made before P3 is generated. The fact that motor response latencies can be varied independently of P3 latencies (24) suggests, in addition, that P3 responding is not a consequence of motor system activity invisible in surface recordings. These considerations are consistent with the idea that motor and P3 systems receive parallel inputs from structures involved in decision processes.

Such functional parallelism would be consistent with a P3 generator located in the limbic system (see reference 26 for a recent discussion of corticolimbic connections). The evident absence of a functional association between P3 and any specific sensory modality or cognitive process (25) also seems consistent with the limbic system's organization and its presumed involvement in motivational processes (26). Moreover, recordings from depth electrodes indicate the presence of P3-like activity in the limbic system (e.g., 27-30).

Responses that resemble P3 can also be recorded in cerebral cortex (31). This is not inconsistent with the notion of a limbic P3 mechanism because limbic efferents project directly or indirectly to widespread regions of cortex (26) where they presumably drive or modulate the responses of their target cells. Other P3-like responses have also been recorded from electrodes in subcortical structures outside the limbic system (32,33). In fact, the number of regions in which P3-like activity can be recorded suggests that, if a single mechanism drives the P3 system, its projections must be very widespread. A system of anatomically distributed projection zones might explain the multiplicity of P3-like responses in surface recordings (34). Whether such zones can be defined, whether their local circuitries explain the response properties of the different P3 waves, and whether their geometries account for the variation in the waves' surface distributions remain to be determined.

A final comment is warranted regarding the generality of the present results. The performance fluctuations examined in this study occurred in the context of four, 10-min blocks of trials. Hence, they probably represent performance variation of the type occurring naturally when individuals engage briefly in relatively uninteresting monitoring tasks. Had our subjects been sleep-deprived or stressed in some other way, quantitatively and qualitatively different phenomena might have been encountered. New ERP components appear, for example, as individuals drift into sleep (35). The properties of these responses are essentially unknown because they have not

been studied in behaving subjects. Nevertheless, these responses also are candidate markers of incipient performance failure.

RECOMMENDATIONS

Future work should be directed toward separating the responses of simultaneously active populations of neurons. These are confounded in surface measurements and are a major source of systematic measurement error. There is no valid way to distinguish between simultaneous responses without knowing their generators' locations. Therefore, it is important to develop a body of theory detailed enough to relate changes in function to changes in anatomically localizable structures.

Examining performance under conditions known to yield high rates of performance failure would be useful. In some of these conditions, ERPs differ qualitatively from ERPs recorded in alert subjects. The period of transition into sleep is an example. Incorporating the peculiarities of those conditions into existing theory should help improve the detection of performance failures.

REFERENCES

1. Trejo, L.J., Lewis, G.W., and Blankenship, M.H., Brain Activity During Tactical Decision-Making: II. Probe-Evoked Potentials and Workload, NPRDC TN 88-12, Navy Personnel Research and Development Center, San Diego, CA, December, 1987.

2. Donchin, E., Kramer, A.F., and Wickens, C.D., "Applications of Brain Event-related Potentials to Problems in Engineering Psychology." In M.G.H. Coles, E. Donchin, and S. Porges (Eds.), Psychophysiology: Systems, Processes, and Applications, Till & Till, Middleton, NJ, pp. 702-718, 1986.

3. Yeh, Y.-Y. and Wickens, C.D., "Dissociation of Performance and Subjective Workload." Human Factors, Vol. 30, pp. 111-120, 1988.

4. Folkard, S., Condon, R., and Herbert, M., "Night Shift Paralysis." Experentia, Vol. 40, pp. 510-512, 1984.

5. Wickens, C.D., "Processing Resources in Attention." In R. Parasuraman and D.R. Davies (Eds.), Varieties of Attention, Academic Press, Inc., Orlando, FL, pp. 63-102, 1984.

6. Allison, T., Wood, C. C., and McCarthy, G. "The Central Nervous System." In M.G.H. Coles, E. Donchin, and S. Porges (Eds.), Psychophysiology: Systems, Processes, and Applications, Till & Till, Middleton, NJ, pp. 5-25, 1986.

7. Hillyard, S. and M. Kutas, "Electrophysiology of Cognitive Processing." Annual Review of Psychology, Vol. 34, pp. 33-61, 1983.

8. Gevins, A.S. and Cutillo, B.A., "Signals of Cognition." In F.H. Lopes da Silva, W. Storm van Leewuen, and A. Remond (Eds.), Clinical Applications of Computer Analysis of EEG and other Neurophysiological Signals, Elsevier Press, New York, NY, pp. 335-384, 1986.

9. Sutton, S., Braren, M., Zubin, J., and John, E.R., "Evoked Potential Correlates of Stimulus Uncertainty." Science, Vol. 150, pp. 1187-1188, 1965.

10. Ruchkin, D.S., Sutton, S. and Mahaffey, D., "Functional Differences Between Members of the P300 Complex: P3e and P3b." Psychophysiology, Vol. 24, pp. 87-103, 1987.

11. Sutton, S. and Ruchkin, D.S., "The Late Positive Complex: Advances and New Problems." In R. Karrer, J. Cohen, and P. Tueting (Eds.), Brain and Information: Event-Related Potentials, Annals of the New York Academy of Sciences, Vol. 425, pp. 1-23, 1984.

12. Ruchkin, D.S. and Sutton, S., "Positive Slow Wave and P300: Association and Disassociation." In A.K.W. Gaillard and W. Ritter (Eds.), Tutorials in ERP Research: Endogenous Components, North Holland, Amsterdam, pp. 233-250, 1983.

13. Bostock, H. and Jarvis, J., "Changes in the Form of the Cerebral Evoked Response Related to the Speed of Simple Reaction Time." Electroencephalography and Clinical Neurophysiology, Vol. 29, pp. 137-145, 1970.

14. Roth, W.T., Ford, J.M., and Kopell, B.S., "Long-latency Evoked Potentials and Reaction Time." Psychophysiology, Vol. 15, pp. 17-23, 1978.

15. Friedman, D., "P300 and Slow Wave: The Effects of Reaction Quartile." Biological Psychology, Vol. 18, pp. 49-71, 1984.

16. Gratton, G., Coles, M.G., and Donchin, E., "A New Method for Off-Line Removal of Ocular Artifact." Electroencephalography and Clinical Neurophysiology, Vol. 55, pp. 468-484, 1983.

17. Davidson, M. and Toporek, J., BMDP 4V--General Univariate and Multivariate Analysis of Variance, BMDP Technical Report No. 67, BMDP Statistical Software, Los Angeles, CA, 1986.

18. Goodin, D.S. and Aminoff, M.J., "The Relationship between the Evoked Potential and Brain Events in Sensory Discrimination and Motor Response." Brain, Vol. 107, pp. 241-251, 1984.

19. Pfefferbaum, A., Ford, J., Johnson, R., Wenegrat, B., and Kopell, B. S., "Manipulation of P3 Latency: Speed vs. Accuracy Instructions." Electroencephalography and Clinical Neurophysiology, Vol. 55, pp. 188-197, 1983.

20. Woody, C.D., "Characteristics of an Adaptive Filter for the Analysis of Variable Latency Neuroelectric Signals." Medical and Biological Engineering, Vol. 5, pp. 539-553, 1967.

21. Fowler, B., Kelso, B., Landolt, J., and Porlier, G., "The Effects of Nitrous Oxide on P300 and Reaction Time." Electroencephalography and Clinical Neurophysiology, Vol. 69, pp. 171-178, 1988.

22. Ritter, W., Simson, R., and Vaughan, H.G. Jr., "Association Cortex Potentials and Reaction Time in Auditory Discrimination." Electroencephalography and Clinical Neurophysiology, Vol. 33, pp. 547-555, 1972.

23. Kutas, M., McCarthy, G., and Donchin, M., "Augmenting Mental Chronometry: The P300 as a Measure of Stimulus Evaluation Time." Science, Vol. 197, pp. 792-795, 1977.

24. McCarthy, G. and Donchin, M., "A Metric for Thought: A Comparison of P300 Latency and Reaction Time." Science, Vol. 211, pp. 77-80, 1981.

25. Fritchard, W., "Psychophysiology of P300." Psychophysiology, Vol. 89, pp. 506-540, 1981.

26. Pandya, D.N. and Yeterian, E.H., "Architecture and Connections of Cortical Association Areas." In A. Peters and E.G. Jones (Eds.), Cerebral Cortex, Vol. 4: Association and Auditory Cortices, Plenum Press, New York, NY, pp. 3-62, 1985.

27. Halgren, E., Squires, N.K., Wilson, C.L., Rohrbaugh, J.W., Babb, T.L., and Crandall, P.H., "Endogenous Potentials Generated in the Human Hippocampal Formation and Amygdala by Infrequent Events." Science, Vol. 210, pp. 803-805, 1980.

28. Altafullah, I., Halgren, E., Stapelton, J.M., and Crandall, P.H., "Interictal Spike-Wave Complexes in the Human Medial Temporal Lobe: Typical Topography and Comparisons with Cognitive Potentials." Electroencephalography and Clinical Neurophysiology, Vol. 63, pp. 503-516, 1986.

29. Wood, C.C., McCarthy, G., Squires, N.K., Vaughan, H.G., Woods, D.L., and McCallum, W.C., "Anatomical and Physiological Substrates of Event Related Potentials: Two Case Studies." Annals of the New York Academy of Sciences, Vol. 425, pp. 681-721, 1984.

30. Stapleton, J.M. and Halgren, E., "Endogenous Potentials Evoked in Simple Cognitive Tasks: Depth Components and Task Correlates." Electroencephalography and Clinical Neurophysiology, Vol. 67, pp. 44-52, 1987.

31. Wood, C.C., and McCarthy, G.A., "A Possible Frontal Lobe Contribution to Scalp P300." In J.W. Rohrbaugh, R. Johnson, Jr., and R. Parasuraman (Eds.), Research Reports: Eighth International Conference on Event-Related Potentials of the Brain (EPIC VIII), Stanford, CA, pp. 164-171, June 1986.

32. Yingling, C.D. and Hosobuchi, Y., "A Subcortical Correlate of P300 in Man." Electroencephalography and Clinical Neurophysiology, Vol. 59, pp. 72-76, 1984.

33. Velasco, M., Velasco, F., Velasco, A.L., Almanza, X., and Olvera, A., "Subcortical Correlates of the P300 Potential Complex in Man." Electroencephalography and Clinical Neurophysiology, Vol. 64, pp. 199-210, 1986.

34. Halgren, E., Stapleton, M.S., Smith, M., and Altafullah, I., "Generators of the Human Scalp P3(s)." In R.Q. Cracco and I. Bodis-Wollner (Eds.), Frontiers of Clinical Neuroscience, Vol. 3: Evoked Potentials, Alan R. Liss, Inc., New York, NY, pp. 269-284, 1986

35. Ujszaszi, J. and Halasz, P., "Long Latency Evoked Potential Components in Human Slow Wave Sleep." Electroencephalography and Clinical Neurophysiology, Vol. 69, pp. 516-622, 1988.

Other Related NAMRL Publications

None are applicable.